AFRL-SR-BL-TR-98-

# REPORT DOCUMENTATION PAGE

| 1. AGENCY USE ONLY (Leave Blank) | 2. REPORT DATE<br>07 May 97 | 3. REPORT TYPE AND DATES COVERED<br>Final (01 Dec 93 – 30 Nov 96) |
|---|---|---|

| 4. TITLE AND SUBTITLE<br>Optimization of Molecular Structures | 5. FUNDING NUMBERS<br>F49620-94-1-0081 |
|---|---|

**6. AUTHORS**
Dr. Teresa L. Head-Gordon

| 7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)<br>Lawrence Berkeley National Laboratory<br>Berkeley, CA 94720 | 8. PERFORMING ORGANIZATION REPORT NUMBER |
|---|---|

| 9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)<br>AFOSR/NM<br>110 Duncan Avenue, Room B-115<br>Bolling Air Force Base, DC 20332-8080 | 10. SPONSORING/MONITORING AGENCY REPORT NUMBER |
|---|---|

**11. SUPPLEMENTARY NOTES**

| 12a. DISTRIBUTION AVAILABILITY STATEMENT<br>Approved for Public Release | 12b. DISTRIBUTION CODE |
|---|---|

**13. ABSTRACT** *(Maximum 200 words)*
There has been improvement in robust prediction of secondary structure of real proteins. New models of aqueous solvation have been used to predict energy surface descriptions for protein structure.

DTIC QUALITY INSPECTED 2

19980129 062

| 14. SUBJECT TERMS<br>robust, proteins, aqueous | | | 15. NUMBER OF PAGES |
|---|---|---|---|
| | | | 16. PRICE CODE |

| 17. SECURITY CLASSIFICATION OF REPORT<br>Unclassified | 18. SECURITY CLASSIFICATION OF THIS PAGE<br>Unclassified | 19. SECURITY CLASSIFICATION OF ABSTRACT<br>Unclassified | 20. LIMITATION OF ABSTRACT<br>UL |
|---|---|---|---|

Progress Report
AFOSR, Optimization of Molecular Structures
September 1, 1996
Revised 5/7/97

## 1. Cover Sheet:

P.I.: Dr. Teresa L. Head-Gordon
Life Sciences Division
Lawrence Berkeley National Laboratory
Berkeley, California 94720
Grant start date: 12/1/93
Grant number: F49620-94-1-0081

## 2. Objectives:

We have new network results that show good improvement in robust
prediction of secondary structure of real proteins, especially in the
category of beta-sheets. We have completed the final testing (known as
jack-kniving) to see that the network can predict well over divisions
of the total database into different training and test sets. We have
submitted a manuscript to report these results.

We have also revised our objectives this year to improve the
energy surface descriptions for protein structure by incorporating new
models of aqueous solvation. This type of improvement is vital for
energetically distinguishing native protein folds from misfolded
proteins. The original gas phase energy surfaces would involve
optimization applied to an objective function that looks like a
golf course. The addition of aqueous solvation, a physically motivated
and more realistic description of proteins and their environment,
qualitatively changes the surface to have a distinct global energy
region corresponding to the native structure, and to correctly
distinguish that region from all misfolded structures which are now
higher in energy. These improved energy surface descriptions will
provide a more realistic energy surface and meaningful application
of global or constrained optimization to the prediction of protein
structure.

## 3. Status of effort:

Protein structure prediction involves two major areas of
research: (1) a realistic energy surface (or objective function), and
(2) a means for rapidly characterizing that energy surface by determining
the global minimum and all low-lying minima on that surface. Our current
status is a major gearing up of my laboratory to use global optimization

methods (stochastic/perturbation from the Schnabel and Byrd groups at
U. of Colorado, Jorge More' at Argonne National Labs) and soft constraint
optimization (defined from my designed neural networks) together to
predict protein structure using high performance computing platforms
such as the Cray T3E. We have also made important progress this year on
a new direction for improving the energy surface, by going from gas phase
function to energy surfaces which can ultimately describe an aqueous
solvent environment, energy functions which are managable in the context
of optimization. This new energy surface will ultimately be combined with
our designed neural networks and global optimization.

## 4. Accomplishments/New Findings:

One new accomplishment this year is the recognition of how to
extract potential of mean force functions that describe aqueous solvation
between amino acid pairs using a combination of neutron scattering
(experimental collaborator: Prof. Robert Glaeser, UC Berkeley) and
molecular dynamics simulations. These wifll be numerically-based
functions and derivatives that will be interfaced with gas phase energy
functions such as AMBER to describe a protein in aqueous environment.
We have realized one publication in press, and two in preparation. The
longer term goal is to use the new energy surfaces in conjunction with
global and constrained optimization. The success and predictive capacity
of global and constrained optimization methods to find low energy or
even global energy structures will rely heavily on a qualitatively
correct energy surface.

Our second accomplishment is the that all pieces of the protein
structure prediction approach are in place: physically derived
constraints from neural networks, global optimization in small subspaces
from the Byrd and Schnabel groups (and future collaborations with Jorge
More' and global continuation), and substantial high performance
computing resources at NERSC.  My group has built a "front end" piece of
code that sets up the energy surface for a given protein for subsequent
global optimization using neural networks and stochastic/perturbation,
has ported this combined piece of software to the Cray T3D and T3E
platforms which exist at the newly established supercomputer center
(NERSC) at LBNL, and starting new improvements on our neural network
designs to predict secondary structure biases that will be used in our
optimization schemes.

## 5. Personnel Supported
Dr. Teresa Head-Gordon (PI)
Richard C. Yu (research assistant)

## 6. Publications:
Published:
[1] R. C. Yu & T. Head-Gordon (1995). Neural-network design
applied to protein-secondary-structure predictions.
Phys. Rev. E. 51, 3619-3627.
[2] A. Pertsemlidis, A. M. Saxena, A. K. Soper, T. Head-Gordon,
R. M. Glaeser (1996). Direct, structural evidence for

modified solvent structure within the hydration shell of
a hydrophobic amino acid. Proc. Natl. Acad. Sci. 93,
10769-10774.

Submitted manuscripts or in preparation
[3] R. C. Yu, I. Dubchak & T. Head-Gordon (1997). High performance
   neural network predictions of protein secondary structure
   without use of sequence or structural homologies.
   Manuscript in preparation.
[4] T. Head-Gordon, J. M. Sorenson, A. Pertsemlidis, R. M. Glaeser
   (1996). Hydration structure near hydrophobic and hydrophilic
   amino acids. Submitted to Biophysical Journal

7. Interactions/Transitions:
   My research has appeared in a publically distributed business
   opportunity manual which encourages CRADAs between industry and
   LBNL in 1995 and 1996 and 1997, and we have copyrighted our
   algorithm. I have had inquiries from industry on using this
   algorithm.

8. New discoveries, inventions, or patent disclosures
   Our patent office determined that software can not be patented,
   and copyrighted it instead.

9. Honors/Awards
   Promotion to computational biology group leader at LBNL
   (do promotions count? maybe not!)